AD________________

AWARD NUMBER DAMD17-96-1-6058

TITLE: Advanced Methods for the Computer-Aided Diagnosis of Lesions in Digital Mammograms

PRINCIPAL INVESTIGATOR: Maryellen L. Giger, Ph.D.

CONTRACTING ORGANIZATION: University of Chicago
Chicago, Illinois 60637

REPORT DATE: July 1998

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1998	Annual (7 Jun 97 - 6 Jun 98)

4. TITLE AND SUBTITLE
Advanced Methods for the Computer-Aided Diagnosis of Lesions in Digital Mammograms

5. FUNDING NUMBERS
DAMD17-96-1-6058

6. AUTHOR(S)
Maryellen L. Giger, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Chicago
Chicago, Illinois 60637

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The objective of the proposed research is to develop computer-aided diagnosis methods for use in mammography in order to increase the diagnostic accuracy of radiologists. Specifically we have developed advanced computerized schemes for the detection spiculated lesions and architectural distortions based on the calculation of the Hough spectrum and for the detection of small, low-contrast early cancers based on gradient and circularity filters. Also, computerized classification schemes for masses using artificial neural networks, rule-based methods, and hybrid systems have been developed. We also investigated a computerized method for including temporal change between mammographic examinations. The efficacy and efficiency of the CAD methods for mammography are being evaluated on a clinical workstation. The potential significance of this research project lies in the fact that if the detectability of cancers can be increased by employing a computer to aid the radiologist's diagnosis, then the treatment of patients with cancer can be initiated earlier and their chance of survival improved.

14. SUBJECT TERMS
Breast Cancer, computer-aided diagnosis, digital mammography, computer vision, artificial intelligence

15. NUMBER OF PAGES
26

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

✓ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Maryellen Geiger 7-1-98
PI - Signature Date

Table of Contents

Page

FRONT COVER.. 1

STANDARD FORM (SF 298).................................... 2

FOREWORD.. 3

INTRODUCTION 5

Nature of the problem

Background of previous work

Purpose of the present work

Methods of approach

BODY 9

Development of advanced computerized schemes for the detection of masses in digital mammograms.

Experimental method
Results to date

Development of advanced computerized schemes for the classification of masses in digital mammograms.

Experimental method
Results to date

Development of computerized methods based on multiple views for enhanced mammographic interpretation.

Experimental method
Results to date

Incorporation of the computer-vision methods with an Mammo/Icon mammographic review system for enhanced diagnosis.

Experimental method
Results to date

Evaluation of the CAD methods for mammography

Experimental method
Results to date

CONCLUSIONS 21

REFERENCES 22

INTRODUCTION

Our first-year report was accepted as an excellent report as submitted. The review of our first year report indicated that the report was well-written with extensive background material and meticulous description of the theoretical basis for the algorithms, which are not necessary in future annual reports and could be referred to with appropriate citations. Thus, in this second-year report, we have substantially shortened the background sections and refer the reviewer to our first-year report.

Nature of the problem

Breast cancer is a leading cause of death in women, causing an estimated 44,000 deaths per year. Mammography is the most effective method for the early detection of breast cancer and it has been shown that periodic screening of asymptomatic women does reduce mortality. Mammography is becoming one of the largest volume x-ray procedures routinely interpreted by radiologists.

Although mammography is currently the best method for the detection of breast cancer, between 10-30% of women who have breast cancer and undergo mammography have negative mammograms. In approximately two-thirds of these false-negative mammograms, the radiologist failed to detect the cancer that was evident retrospectively. Low conspicuity of the lesion, eye fatigue and inattentiveness are possible causes for these misses. We believe that the effectiveness (early detection) and efficiency (rapid diagnosis) of screening procedures could be increased substantially by use of a computer system that successfully aids the radiologist by indicating locations of suspicious abnormalities in mammograms.

In addition, many breast cancers are detected and referred for surgical biopsy on the basis of a radiographically detected mass lesion or cluster of microcalcifications. Although general rules for the differentiation between benign and malignant breast lesions exist, considerable misclassification of lesions occurs with the current methods. On average, only 10-30% of masses referred for surgical breast biopsy are actually malignant. Surgical biopsy is an invasive technique that is an expensive and traumatic experience for the patient and leaves physical scars that may hinder later diagnoses (to the extent of requiring repeat biopsies for a radiographic tumor-simulating scar). A computerized method capable of detecting and analyzing the characteristics of benign and malignant masses, in an objective manner, should aid radiologists by reducing the numbers of false-positive diagnoses of malignancies, thereby decreasing patient morbidity as well as the number of surgical biopsies performed and their associated complications.

The development of computer methods to assist radiologists is a timely project in the sense that digital radiography is on the threshold of widespread clinical use. The arrival of digital radiographic systems allows for the acquisition of image data in a format accessible to computerized schemes. The potential significance of this research project lies in the fact that if the detectability of cancers can be increased by employing a computer to aid the radiologist's diagnosis, then the treatment of patients with cancer can be initiated earlier and their chance of survival improved.

Background of previous work

Breast cancer is a leading cause of death in women, causing an estimated 44,000 deaths per year (1). Mammography is the most effective method for the early detection of breast cancer (2-5) and it has been shown that periodic screening of asymptomatic women does reduce mortality (6-11). Various medical organizations have recommended the use of mammographic screening for the early detection of breast cancer (3). Thus, mammography is becoming one of the largest volume x-ray procedures routinely interpreted by radiologists.

It has been reported that between 30 to 50% of breast carcinomas detected mammographically demonstrate clusters of microcalcifications (12-14), although about 80% of breast carcinomas reveal microcalcifications upon microscopic examination (15-18). In addition, studies indicate that 26% of nonpalpable cancers present mammographically as a mass while 18% present both with a mass and microcalcifications (19). Although mammography is currently the best method for the detection of breast cancer, between 10-30% of women who have breast cancer and undergo mammography have negative mammograms (20-24). In approximately two-thirds of these false-negative mammograms, the radiologist failed to detect the cancer that was evident retrospectively (23-26). Low conspicuity of the lesion, eye fatigue and inattentiveness are possible causes for these misses. It has been suggested that double reading (by two radiologists) may increase sensitivity (27-29).We believe that the effectiveness (early detection) and efficiency (rapid diagnosis) of screening procedures could be increased substantially by use of a computer system that successfully aids the radiologist by indicating locations of suspicious abnormalities in mammograms.

Many breast cancers are detected and referred for surgical biopsy on the basis of a radiographically detected mass lesion or cluster of microcalcifications. Although general rules for the differentiation between benign and malignant breast lesions exist (20,30), considerable misclassification of lesions occurs with the current methods. On average, only 10-30% of masses referred for surgical breast biopsy are actually malignant (20,31). Surgical biopsy is an invasive technique that is an expensive and traumatic experience for the patient and leaves physical scars that may hinder later diagnoses (to the extent of requiring repeat biopsies for a radiographic tumor-simulating scar). A computerized method capable of detecting and analyzing the characteristics of benign and malignant masses, in an objective manner, should aid radiologists by reducing the numbers of false-positive diagnoses of malignancies, thereby decreasing patient morbidity as well as the number of surgical biopsies performed and their associated complications.

The development of computer methods to assist radiologists is a timely project in the sense that digital radiography is on the threshold of widespread clinical use. The arrival of digital radiographic systems allows for the acquisition of image data in a format accessible to computerized schemes. The potential significance of this research project lies in the fact that if the detectability of cancers can be increased by employing a computer to aid the radiologist's diagnosis, then the treatment of patients with cancer can be initiated earlier and their chance of survival improved.

Work of others in the field

Comprehensive summaries of investigations in the field of mammography CAD have been published by the P.I. (30,31). In the 1960's and 70's, several investigators attempted to analyze mammographic abnormalities with computers (34-42) and these are described in our first annual report. These previous studies demonstrated the potential capability of using a computer in the detection of mammographic abnormalities.

Computer-aided diagnosis, in general, has attracted little attention twenty years ago, perhaps due to the inconvenience involved in obtaining a radiograph in digital format. Recent work, though, shows a promising future for both the computerized detection and diagnosis of mammographic lesions (43-58) and these methods are also described in our first annual report.

Previous relevant work of the investigators

We in the Kurt Rossmann Laboratories for Radiologic Image Research at The University of Chicago have vast experience in developing various computer-aided diagnosis (CAD) methods in mammography, chest radiography, and angiography (59-86). Our CAD methods in digital mammography, which include the computerized detection of microcalcifications and masses, have

achieved useful levels of sensitivity and specificity and thus are currently undergoing clinical testing. As suggested, the descriptions have been shortened here (but can be found in our first annual report).

Computerized detection of clustered microcalcifications

Our detection scheme for clustered microcalcifications includes a preprocessing step referred to as a difference-image approach (59,60). Basically, the original digital mammogram is spatially filtered twice: once to enhance the signal-to-noise ratios of the microcalcifications and a second time to suppress them. The difference between the two resulting processed images yields an image (a difference image) in which the variations in background density are largely removed. Microcalcifications are then segmented from the difference image using global gray-level thresholding and local thresholding techniques. The segmented image is next subjected to feature-extraction techniques in order to remove signals that likely arise from structures other than microcalcifications, including size analysis, texture analysis, and clustering analysis (62-65,76,85). In addition, a shift-invariant neural networks is used directly on the digital image data in order to enhance microcalcifications and reduce false-positive detections.

Computerized detection of masses

Our initial method for the detection of mammographic masses is based on deviations from the architectural symmetry of normal right and left breasts, with asymmetries indicating potential masses (66-68,74). The input to the computerized scheme, for a given patient, are the four conventional mammograms obtained in a routine screening examination: the right cranio-caudal (CC) view, the left CC view, the right medio-lateral-oblique (MLO) view, and the left MLO view. After automatic registration of corresponding left and right breast images, a nonlinear subtraction technique is employed which includes a run-length analysis to yield two images that contain locations of suspected masses for the left and right breasts (75). Next, feature-extraction techniques, which include morphological filtering and analysis of size, shape and distance from border, are used to reduce the number of false-positive detections. These features are merged using an artificial neural network.

Use of artificial neural networks for detection and classification

We have also investigated the application of artificial neural networks to the detection and classification of mammographic lesions (65, 69).

Pre-clinical testing of an "intelligent" mammography workstation

We have implemented our computerized detection schemes for masses and clustered microcalcifications on a prototype "intelligent" mammography workstation in an ongoing clinical study (70,72,85). The workstation, on which the automated computerized detection schemes are run, consists of a film digitizer, a high-speed computer, a magneto-optical jukebox, and hard and soft copy displays. Preliminary prospective evaluations of the prototype, which was installed in the clinical mammography reading area at the University of Chicago Medical Center on November 8, 1994, are promising.

Computerized classification of mammographic masses

Malignant masses often can be distinguished from benign masses due to their more spiculated appearance in the mammographic image. Thus, in the classification of masses, our computerized scheme is based on the degree of spiculation exhibited by the mass in question (77,82). With a pathologically-confirmed database of 95 masses (57 malignant and 38 benign), the classification scheme achieved an A_Z (area under the ROC curve) of 0.90 using a rule-based scheme together with a conventional feed-forward back-propagation neural network.

The Mammo/Icon system

Dr. Swett's group began investigating methods for strengthening radiologic diagnosis in conjunction with Dr. Perry Miller. They developed an expert system called Mammo/ICON which was able to generate context sensitive English prose critiques containing advice about radiologic diagnosis (57, 58). This system also supplements critiques with display of reference images selected to confront issues raised by a given clinical case. The system is used to report mammographic findings, using either speech recognition technology, or mouse interaction with a graphical user interface. As findings are recorded, Mammo/Icon functions in the background, collecting potentially helpful images from a knowledge base.

Purpose of the present work

The main hypothesis to be tested is that given dedicated computer-vision programs for the computer-assisted interpretation of mammograms, the diagnostic accuracy for mammographic interpretation will be improved, yielding earlier detection of breast cancer (i.e., a reduction in the number of missed lesions) and a reduction in the number of benign cases sent to biopsy. Computer-aided diagnosis (CAD) can be defined as a diagnosis made by a radiologist who takes into consideration the results of a computerized analysis of radiographic images and uses them as a "second opinion" in detecting lesions and in making diagnostic decisions. The final diagnosis would be made by the radiologist.

Methods of approach

The objective of the proposed research is to develop computer-aided diagnosis methods for use in mammography in order to increase the diagnostic decision accuracy of radiologists and to aid in mammographic screening programs. The CAD methods will include a parallel method for the detection of a range of mass types and for the incorporation of information from multiple views (i.e., CC and MLO, and prior mammograms).

The **specific objectives** of the research to be addressed are:

(1) Development of advanced computerized schemes for the detection and classification of masses in digital mammograms.

(a) Development of a computerized detection scheme for spiculated lesions and architectural distortions based on the calculation of the Hough spectrum.

(b) Development of a computerized detection scheme for small, low-contrast early cancers based on gradient and circularity filters.

(c) Incorporation of the two new methods with a previously-developed bilateral-subtraction method along with feature analyses into a system for lesion detection.

(d) Further development of computerized classification schemes for masses.

(2) Development of computerized methods based on multiple views for enhanced mammographic interpretation.

(a) Development of computerized methods for the incorporation of image information from the CC and MLO views of mammographic examinations.

(b) Development of computerized methods for analysis of temporal change between mammographic examinations.

(3) Incorporation of the computer-vision methods with an Mammo/Icon mammographic review system for enhanced diagnosis.

(a) Expansion of the Mammo/Icon database descriptors to include CAD derived parameters.

(b) Calculation of the computer extracted features of images in the Mammo/Icon database.

(c) Development of hardware and software interfaces for CAD and Mammo/Icon.

(4) Evaluation of the CAD methods for mammography.

BODY: Experimental methods and results to date

Development of advanced computerized schemes for the detection of masses in digital mammograms.

Experimental methods

Additional detection methods for masses are needed due to the variation in types of masses. Our current bilateral-subtraction method has been shown to be successful in the detection masses that are apparent from the deviation from symmetry between the left and right breasts. However, very spiculated lesions and architectural distortions are not always detected with the bilateral subtraction method and very small, low-contrast cancer lesions also are difficult to detect using the bilateral subtraction method.

For each digital mammogram, prior to the computer analysis for the detection of lesions, it is necessary to identify the breast region from the rest of the film regions. The breast region will be automatically segmented by excluding uniform dark (direct exposure) and uniform bright (unexposed) film regions. Initial noise filtering using a median filter will be applied to the digital image followed by application of a gray-value range operator (80). Using information from the local range operator a modified global histogram analysis will be performed. Region growing will be performed on the threshold image, followed by a morphological erosion operation. A distance map of the image will be determined and the boundary of the segmented object in the image will be then tracked to yield its contour. The contour will then be used in the subsequent image analysis schemes. Our initial evaluation of this segmentation algorithm involved the analysis of 740 digital mammograms with a resulting 97% of the detected contours acceptable for CAD usage (79).

We are developing methods for the segmentation of dense portions in the breast since such regions can obscure the visibility of lesions and distortions. Determination of the dense portions within the breast region is done using gray-level histogram analysis within the breast region as well as within a small region along the chest wall. This selection of ROI location is due to the fact that as one approaches the chest wall within a mammogram, the amount of dense tissue decreases, and thus an estimation of the pixel value corresponding to fatty tissue can be calculated. The ROI, of predetermined width and height is positioned at a predetermined distance (mm or pixels) from the chest wall. The location of the chest wall side in the mammogram is determined during the breast segmentation step since it differs greatly from the external side of the breast. By analysis of the bimodal nature of the histograms, a peak will be located corresponding to the dense region cutoff. Gray-level thresholding within the breast region is then performed. Thus, the ROI near the chest wall, as compared to that of the entire breast regions can be used to indicate the gray levels of the fatty portions. With a database of 700 mammograms, the dense portions as segmented by the computer and those segmented by a radiologist had good correlation at a statistical significant level.

Mammographic spiculated lesions and architectural distortions are usually associated with malignancy, which makes them important signs in the screening of breast cancer. The Hough spectrum technique is developed from the traditional Hough transform, which is effective in the description of geometric structures. For our investigation, Hough spectra will be calculated for regions-of-interest (ROIs) within the breast region of the digitized mammogram (84). Thresholding will then be conducted with the threshold level based on the statistical properties of the spectra. Those ROIs with strong signals of spiculation or architectural distortion will be indicated as regions of potential lesions. We will also investigate the effect of ROI size on performance. Details on the Hough spectrum technique can be found in our first annual report.

The radiographic appearance of the breast tissue is abundant in the textural information composed of bright and slender unite structures. Generally, the distribution of these textural element takes the pattern of radiating from nipple to the chest wall, which makes them roughly parallel to one another

locally. The presence of certain abnormal structures such as a spiculated mass or an architectural distortion, however, may alter this trend by generating another radiating center, thus changes the textural appearance dramatically in a neighboring area. The basic structural unit of the mammographic textural pattern can be modeled as stripe. Hence, the technique can be applied to analyze the mammographic textural pattern, especially for the detection of the above mentioned abnormal structures. In the Hough-spectrum-based analysis, ROIs are placed within the breast region of the digitized mammogram and from the image data within each ROI a Hough spectrum is calculated (87).

In addition, with screening mammography, an increasing number of small invasive breast cancers is found, which can be seen on mammograms as small circumscribed mass often < 1 cm. Current computer-aided detection schemes based on bilateral subtraction are not well suited for these small masses, due to significant normal variations between the right and left breast . The purpose of this task is to develop a single-image method specifically for the computer-aided detection of small circumscribed masses. The method, we are developing, is based on a modified median filter, a modified morphological open operation, filtering with a mass filter for the initial detection of circumscribed densities, matching using a deformable shape template with Fourier descriptors, characterization of the match using simulated annealing, and measuring the circularity and density characteristics of the suspected lesion. After a 3x3 median filtering step, a morphological open (erosion followed by dilation) operation with a 7 pixel wide circle as structuring element is used to eliminate small circular and thin linear structures below a certain diameter (83). To preserve the gray value characteristics of larger lesions as far as possible, only pixels with a small difference to the local minimum are used as erosion centers. If the gray value after dilation exceeds the original pixel value, the original pixel value is used instead of the filtered value.

A newly designed second derivative filter with a circular, 21-pixel wide base identifies circumscribed density peaks in the image. The filter value is based on the local gradient (7x7 kernel) in x- (D_x) and y- (D_y) direction and the filter value is calculated separately for 16 edge orientation bins γ. The final filter value is then calculated as the sum of the individual orientation bins with omission of the 4 bins with the highest values:This prevents an influence of straight edges (e.g. the pectoralis muscle border) on the filter value without changing the filter value for ideal circular lesions. Local maxima of the filter value define potential center positions of mass lesions, which are used in the next step, the matching of a deformable template onto the lesion border. In this process, a deformable shape template is created by inverse Fourier transform of a limited number of complex Fourier terms. The final lesion contour is then identified by variation of the Fourier terms within a certain range with minimization of a cost function based on lesion contrast and edge strength using simulated annealing. For further characterization, a rectangular ROI containing the lesion is extracted from the original peripheral density corrected image. This is used to calculate lesion size, lesion contrast and a radial gradient index RGI. The RGI is a measure of the circularity and density characteristics of a lesion and approaches 1.0 for a ideal circular lesion. It is used to differentiate between true and false positive detections. The algorithm is used repetitively at different resolutions with the pixel size varying from 0.5 mm to 4 mm. Each resolution step covers a certain range of lesion sizes. In the final step, the different detected lesions are integrated into one final result. In case of overlap between different detected lesions, the lesion with the smaller radial gradient index is eliminated.

Output from the three detection preprocessing methods (i.e., the bilateral-subtraction technique, the Hough-spectrum-based technique and the gradient/circularity filter technique) will be merged. The output from each method will go through its own feature analysis method in order to reduce false positives. We currently have 90 features that can be calculated for a suspect lesion both at high resolution (0.1 mm pixel size) and at low resolution (0.5 mm pixel size). We will determine which features are appropriate for each of the three preprocessing methods. These features will be selected based either on one-dimensional analysis (86) or on multi-dimensional analysis involving genetic algorithms (89). False positive reduction will occur based solely on the type of preprocessing methods which will enhance the abnormality. Further reduction in false-positive detections will occur

during the feature-analysis stage. In the feature-analysis stages, regions deemed suspicious by the method will be subjected to geometric-based measures, intensity-based measures, edge-gradient-orientation methods and texture-based measures. (82,86,89). We have successfully developed and used these various types of features to reduce false-positive detections in our other CAD schemes. We realize, of course, that features and the combination of such features used for reducing false positives in mass detections schemes will differ in the proposed scheme for spiculated lesions and architectural distortions. Features used will depend on the specific type of false positives generated by the specific method. It should be noted that false positives generated by the Hough-spectrum-based method may differ from those generated by the bilateral subtraction method and by the gradient/circularity method.

We will evaluate artificial neural networks (ANN) as a means to merge the various features obtained from the computer analysis of the mammograms [in a manner we have successfully used in the past for merging mammographic feature data (69)]. The various features will serve as input data and will be supplied to the input units of the neural network. Prior to input to the ANN, the features will be normalized between 0 and 1. The output data from the neural network are then obtained through successive nonlinear calculations in the hidden and output layers. The calculation at each unit in a layer includes a weighted summation of all entry numbers, an addition of a certain offset number, and a conversion into a number ranging from 0 to 1 using a sigmoid-shape function such as a logistic function. The neural network will be trained by a back-propagation algorithm using pairs of training input data and desired output data. The desired output data will be initially 1 if the suspect lesion is an actual lesion and 0 otherwise. Once trained, the neural network will accept features and output a value will be related to the likelihood of being an actual lesion (i.e., a true-positive). Feature selection will be performed by analyzing the average and standard deviation of the various features for both high and low risk subjects. Az values will be calculated for each feature as well as for the output of the ANNs. In addition, genetic algorithms, which we have used, in a pilot study, for optimizing feature selection for the task of distinguishing true-positive and false-positive mass detections, will also be used (89).

In the parallel system, images will be analyzed by all three computerized detection methods and the output from each will be merged in a single result file. During grouping either the ANN output value from each detection method will be input to a "grouping" ANN or a logical OR operation followed by a distance grouping operation will be performed. The "grouping" ANN will take as one input a zero value if, for example, only two of the three detection methods identify a suspicious region. In the logical OR operation the sensitivity will increase, but so will the false-positive rate, and thus, the subsequent grouping operation is necessary. In the grouping operation, if two detections are within a specific distance from each other, than the locations will be merged to yield a single identification. The training of these operations will be performed as discussed for other ANNs above. We currently use a grouping operation for the output of the bilateral subtraction method.

Results to date

With the single-image method for detection of small invasive breast cancers (83) localized density peaks on mammograms are identified using a gradient/circularity filter. Lesion contours were generated by matching a deformable template onto a second derivative edge map. In a preliminary study (without further feature analyses to reduce false positives) using 45 non-palpable invasive breast cancers, all with a size less than 1 cm (median size of 7 mm), 82% of the cancers were detected with an average false-positive rate of 2.8 per image (104).

In the Hough spectrum geometric texture analysis technique, the mammogram is analyzed ROI by ROI (84). Each ROI is transformed into its Hough spectrum and then thresholding is performed with its threshold level based on the statistical properties of the spectrum. ROIs with strong signals of spiculation are then screened out as regions of potential lesions. In a preliminary study, 32 images containing spiculated lesions/architectural distortions (biopsy confirmed) were analyzed using

information extracted from the Hough spectrum. Our preliminary studies, using only the Hough spectrum based technique without further feature analyses to reduce false positives, yielded sensitivities of 81% for spiculated masses and 67% for architectural distortions at false positives rates of 0.97 and 2.2 per image, respectively (105). We have also converting the method into an AVS based program to expedite the development and optimization of the parameters such as ROI size.

Output from the bilateral subtraction method and that of the gradient/circularity filtering were combined and analyzed. Many masses were detected by both preprocessing methods. For a database of 20 cancer cases, the bilateral yielded a sensitivity of 75% (at 1.8 false-positives detections per image) and the gradient/circularity filter yielded a sensitivity of 70% at the same false postivie rate. Upon comparison, the gradient/circularity filter found masses that the bilateral did not, thus allowing the sensitivity to increase to 80%. We are currently comparing the false positive overlap to determine the false-positive rate for the combined scheme as well as give us a means to improve the false-positive rate while optimizing the sensitivity for each method.

Since November 8, 1994, all screening mammograms taken at the University of Chicago Hospitals have been analyzed on our clinical prototype mammography worstation, except during downtimes. Downtime has been minimal, less than 20 days in total, which includes a 3-week period when the mammography section moved to a new outpatient center. During that move, networking problems in the new facility contributed to computer system difficulties. For cases in which a cancer was detected, we also retrospectively reviewed any previous mammograms that were in our study cohort. Two radiologists independently reviewed the cases and stated whether the cancer was visible in a previous exam and whether, knowing that the lesion was present, that they would call the patient back for a diagnostic exam based on the findings in the previous exam. In this way, the number of cancers detected by the computer that were initially missed by the radiologists was determined. As of May, 1, 1998, over 14,000 cases have been analyzed (106). With follow-up on the first 10,000 cases, 61 patients have been diagnosed with breast cancer. In 12 of these cases, the screening mammogram(s) were negative even in retrospect. For the mammographically visible cases (n=49), the sensitivity of the two schemes was 68% (34/49). Clinically, 96% of the cancers were detected (47/49). More important than the absolute sensitivity of the workstation is its ability to detect breast cancers that may be missed by a radiologist. In 30 of the 61 cancers, the patient had a screening exam that was read as negative and was included in our study. That is, a screening mammogram that was read as normal, which preceded the cancer being diagnosed. In 14 of these cases, no lesion could be seen in retrospect, i.e., mammographically negative. In 9 of 16 cases, the computer was able to identify the region on the negative-read (cancer visible in retrospect) screening mammogram that corresponded to where the cancer was subsequently detected. Overall, the computer was able to identify the cancer approximately one year before it was diagnosed in approximately 15% (9/61) of all cancer cases and in 56% (9/16) of cases were the cancer was visible in retrospect on a negative-read screening mammogram. The false-positive rate was approximately 1.3 false clusters per image and 2.1 false masses per image. The types of false-positive detections found by the computer in mass detection and clustered microcalcification detection were investigated for 1296 cases. Of the false positives that were indicated by the computer, over 80% of the mass false positives were due to nodular densities on the film.

In order to determine the effect of false-positive detections on mammographic interpretation, we calculated the call-back rate in one-year time periods before and after implementation of the workstation in the clinical area. The callback rate is the fraction of screening mammograms read as abnormal. Before introduction of CAD, 13.2% of screeners were called back for further workup and after the introduction of CAD, 12.6% of screeners were called back for further workup. Thus, the false-positive output from the computer did not increase the number of women called back.

A new development, which is now being implemented into the various detection and classification schemes for mammographic masses, is a new region growing algorithm (107). The segmentation of lesions from surrounding background is a vital step in many computerized mass

detection schemes. We have developed two novel lesion segmentation techniques -- one based on a single feature called the radial gradient index (similar feature to that described above) and one based on a simple probabilistic model to segment mass lesions from surrounding background. In both methods a series of image partitions is created using gray-level information as well as prior knowledge of the shape of typical mass lesions. With the former method the partition that maximizes the radial gradient index is selected. In the latter method, probability distributions for gray-levels inside and outside the partitions are estimated, and subsequently used to determine the probability that the image occurred for each given partition. The partition that maximizes this probability is selected as the final lesion partition (contour). We tested these methods against our previous region-growing algorithm using a database of biopsy-proven, malignant lesions and found that the new lesion segmentation algorithms more closely match radiologists' outlines of these lesions. At an overlap threshold of 0.30, gray level region growing correctly delineates 62% of the lesions in our database while the radial gradient index algorithm and the probabilistic segmentation algorithm correctly segment 92% and 96% of the lesions, respectively. With these new segmentation results we hope to find and extract new features that will help differential between actual lesions and false-positive detections, thus improving the overall performance of computerized mass detection.

Development of advanced computerized schemes for the classification of masses in digital mammograms.

Experimental methods

Spiculation is a primary sign of malignancy for masses detected by mammography. We have developed a technique that analyses patterns and quantifies the degree of spiculation present (82). Our current approach involves (1) automatic lesion extraction using region growing and (2) feature extraction using radial edge-gradient analysis. Two spiculation measures are obtained from an analysis of radial edge-gradients. These measures are evaluated in four different neighborhoods about the extracted mammographic mass. The performance of each of the two measures of spiculation was tested on a database of 95 mammographic masses using ROC analysis that evaluates their individual ability to determine the likelihood of malignancy of a mass. The dependence of the performance of these measures on the choice of neighborhood was analyzed. We have found that it is only necessary to accurately extract the approximate outlines of a mass lesion for the purposes of this analysis since the choice of a neighborhood that accommodates the thin spicules at the margin allows for the assessment of margin spiculation with the radial edge-gradient analysis technique. Two promising measures are the FWHM and the average radial gradient which correspond to the degree of spiculation and how ill-defined is the margin, respectively. The two measures performed at their highest level when the surrounding periphery of the extracted region is used for feature extraction, yielding A_z values of 0.83 and 0.85, respectively, for the determination of malignancy. These are similar to that achieved when a radiologist's ratings of spiculation (A_z=0.85) are used alone. The maximum value of one of the two spiculation measures (FWHM) from the four neighborhoods yielded an A_z of 0.88 in the classification of mammographic mass lesions.

In this objective, we plan to use artificial neural networks along with other measures of the mass in question to obtain an estimate of the likelihood of malignancy. The margin, shape and density of a mass are the three major characteristics used by radiologists to classify mass lesions. Among the three, margin characteristics of a mass are considered to be the most important indicators of its benign and malignant status. The margin of a mass can be categorized as circumscribed, microlobulated, obscured, indistinct or spiculated with a spiculated margin being the strongest sign of malignancy. After the margin of a mass is accurately identified, seven features related to the margin, shape and density of a mass will be extracted from the neighborhoods of the identified mass region: three of them provide margin information; one is used to describe the shape of the mass; the rest are extracted to estimate the density of the mass. Margin information will be obtained.

The shape of a mass can be described as irregular, lobulated, round or oval. It is difficult to accurately quantify the irregularity of a mass. The oval or round shape of a mass will be determined by a elongation measure, which is defined as the ratio of the long axis to the short axis of the mass. In order to accurately determine the shape with the elongation measure, computer-extracted margin will be first smoothed with a morphological open filter. The elongation measures of a mass is calculated from the best fitted round/oval shape based on the smoothed margin of the mass. Although singularly, anatomic density of a mass may not be as powerful as margin or shape related features in distinguishing between benign and malignant masses, taken with the these features, the density assessment can be extremely useful For example, the evaluation of density of a mass is of great importance in diagnosing the masses in the non-spiculated category -- circumscribed, lobulated, indistinct, and obscured. Because the density of a mass is very difficult to accurately access, we introduce three density related measures to estimate the density of a mass lesion from different aspects, which is similar to the ones used by radiologists. They are the average gray-level of a mass, the contrast --the gray-level difference between the mass and its surrounding, and a texture measure -- standard deviation of the average gradient within a mass, which quantifies texture patterns such as veins, trabecular, or other structures that can be seen "through" from a low-density mass, but not a high density mass. A low-density mass tends to have a high value of the texture measure, and low values of average gray-level and contrast, whereas, a high-density mass tends to have a low value of the texture measure; and high values of average gray-level and contrast.

Results to date

We are investigating the potential usefulness of computer-aided diagnosis as an aid to radiologists in the characterization and classification of mass lesions in mammography. Ninety-five mammograms containing masses from 65 patients were digitized. Various features related to the margin, shape and density of each mass were extracted automatically from the neighborhoods of the computer-identified mass regions. Selected features were merged into an estimated likelihood of malignancy using three different automated classifiers. The performance of the three classifiers in distinguishing between benign and malignant masses was evaluated by receiver operating characteristic (ROC) analysis, and compared with those of an experienced mammographer and of five less experienced mammographers (108). Our computer classification scheme yielded an A_Z value of 0.94, similar to that of an experienced mammographer (A_Z=0.91) and statistically significantly higher than the average performance of the radiologists with less mammographic experience (A_Z=0.80). With the database we used, the computer scheme achieved, at 100% sensitivity, a positive predictive value of 83%, which was 12% higher than that of the experienced mammographer and 21% higher than that of the average performance of the less experienced mammographers at a *p*-value of less than 0.001 (108). Thus, automated computerized classification schemes may be useful in helping radiologists distinguish between benign and malignant masses.

We have also investigated the effect of dominant features on neural network performance in the task of classification of mammographic lesions (109). Two different classifiers, an artificial neural network (ANN) and a hybrid system (one step rule-based method followed by an artificial neural network) were investigated to merge computer-extracted features in the classification of malignant and benign masses. Four computer-extracted features were used in the study: spiculation, margin sharpness and two density-related measures. ROC analysis showed that the hybrid system performed significantly better than the ANN method at the high sensitivity levels, yielding an A_z of 0.94 with a specificity of 69% at 100% sensitivity, whereas, the ANN method yielded an A_z of 0.90 with a specificity of 19% at 100% sensitivity. To understand the difference between the two classifiers in their performance, we investigated their learning and decision-making processes by studying the relationships between the outputs and input features. The correlation study showed that the outputs from the ANN alone method strongly correlated with one of the input features (spiculation measure), yielding a correlation coefficient of 0.91 while the correlation coefficients (absolute value) for the other features range from 0.19 to 0.40. The strong correlation between the ANN output and spiculation

measure indicates the learning and decision-making processes of the ANN alone method was dominated by the spiculation measure. A series of three-dimensional plots of the computer output as functions of the input features demonstrate that the ANN method did not learn as effectively as the hybrid system from the other three features in differentiating subtle (non-spiculated) malignant masses from benign masses, thus, resulting in the inferior performance at the high sensitive levels. We found that with a limited database, it is detrimental for an ANN to learn the significance of other features in the presence of a dominant feature. The hybrid system, which initially applied a rule on the spiculation measure prior to an ANN, prevents the over-learning from the dominant feature and performed better than the ANN alone method in merging the computer-extracted features into a correct diagnosis on the malignancy of the masses.

Currently in mammography, the digital image on which CAD analysis is performed is obtained by digitizing a screen-film mammogram. Since the image is sampled when digitized, the digitization of a image using two different scanners will not produce exactly the same digital image (because of different designs, sampling aperture, sampling distance and internal electronic noise, etc. of the laser scanners and the different calibration curves for the transformation of the optical density (OD) to pixel value). Thus, the contrast, noise and resolution of the two images may differ. Thus, as long as CAD analysis relies upon digitized screen-film images, a CAD system (film digitization and computer analysis) may suffer from the variability in the digital formats of a image, which may lead to variations in the performance of the CAD scheme. Two different databases and three different digitizers were involved in this study. One database consisted of 95 mammograms collected from 65 cases: 39 biopsy-confirmed malignant cases, 25 biopsy-confirmed benign cases and one benign case which was determined through more than five years of follow-up. These mammograms were digitized using an optical drum scanner (FIP II, Fuji Film, Tokyo, Japan) at a sampling distance of 0.1 mm and 10-bit quantization. Another database consisted of 110 new cases which were collected from the University of Chicago Radiology files. Of these, 50 cases are biopsy-confirmed malignant, 50 cases are biopsy-confirmed benign diseases and 10 cases are aspiration-confirmed cysts. For each case, two standard views of the affected breast were chosen from a single screening exam. Of the 110 cases, 8 cases had a mass appearing on one view only. Each mammogram this second database was digitized twice using two different laser scanners -- a Konica digitizer (LD 4500; Konica Medical, Wayne, NJ) at 0.1-mm pixel size and 10-bit quantization and a Lumisys laser scanner (Lumiscan 100, Lumisys, Sunnyvale, CA) at a 0.1-mm pixel size and 12-bit quantization. In the evaluation of our classification scheme, both A_z and ${}_{0.90}A_z'$ are important indices. The A_z value was used to evaluate the overall performance, while the partial area index (${}_{0.90}A_z'$) was designed to evaluate the performance of a scheme at a preselected high sensitivity level for those who are interested in knowing the performance at the high sensitivity (110). In addition, the difference in the partial area index ${}_{0.90}A_z'$ quantitatively evaluates, to some degree, the difference in the shape of the two ROC curves. The differences in A_z between the two digital formats were the same for both the ANN-alone and hybrid classifiers. Two-tailed p values obtained from the CLABROC (111) programs showed that the difference in the performance of the classification scheme, due to the difference between the two digitization techniques, using both the ANN and the hybrid classifier were not statistically significant at the level of 0.05 in terms of the A_z and ${}_{0.90}A_z'$.

Development of computerized methods based on multiple views for enhanced mammographic interpretation.

Experimental methods

It is common for radiologists to utilize the correlation of certain lesions in two views in order to verify difficult and ambiguous cases and also to eliminate some possible false positive findings. We plan to investigate the relationship between the locations of lesions (and other landmarks) in breast images obtained with the CC and MLO views. If a suspicious region in one view is detected by our CAD scheme, this relationship will be used to indicate the range of the potential locations of the

corresponding region in another view. Since the variation of a location of a lesion in a breast can be very complex in two different views, we plan to employ an artificial network to learn the relationship between locations of lesions in two different views. This relationship will be used to eliminate some false positive findings and also to verify "true" lesions.

Initially each breast will be scaled to a "standard" breast size prior to coordinate determination and input to an ANN. This will be accomplished by scaling based on the distance from the nipple to the chest wall, and by fitting using the matching method employed in our bilateral-subtraction method. A simple way of determining the coordinates of a lesion employs the Cartesian coordinates of the lesion relative to the film edges. However, since the location of a breast image can be shifted easily by variation in the positioning of the breast in a mammography unit, we plan to investigate a more accurate method of determining the coordinates of the lesion in mammograms. We will employ a polar coordinate system based on the nipple position and the chest wall. For a CC view, the origin of the polar coordinates will be at the nipple position. The distance to the lesion will be measured from the nipple, and the angle will be determined from the line drawn from the nipple to the chest wall. For an MLO view, the origin of the polar coordinates will again be at the nipple position, and the angle will be determined from the line drawn from the nipple to the pectoral muscle.

A three-layer, feed-forward neural network will be used to correlate locations of the same lesion or the same landmarks in two different views. For each set of breast images, two neural networks will be used, namely, (1) a neural network for determining the range of the locations in MLO view when a lesion is found in CC view, and (2) another neural network for determining the range of locations in CC view when a lesion is found in MLO view. During training of the first neural network, the locations of lesions or landmarks in CC view will be entered to the first neural network's input units in terms of their coordinates, and the corresponding locations in MLO view will be given to the output units. The numbers of input units and output units will be equal to or greater than the total number of potential locations of lesions in all mammograms within a given breast size category in the respective view. For the second neural network, the locations in the MLO and CC views will be provided to units in the input and output layers, respectively.

For training each neural network, a unit in the input layer corresponding to a location of a lesion in polar coordinates in one view will have an input value of 1.0, and all other input units with a value of 0. A unit in the output layer corresponding to the location of the same lesions in another view will be given a desirable (target) output value of 1.0, and all other output units with a value of 0. We plan to use a "jack-knife" method for evaluating the performance of the trained neural network. With this method, a randomly-selected half of the data set will be used for training, and the other half will be used for testing. Results will be evaluated by ROC analysis, and the area under the ROC curve will be used as a measure of performance. We have used these methods successfully in the past for evaluating CAD and ANN. A number of parameters related to the neural networks, such as the number of hidden units and the number of iterations required for training, will be determined empirically based on ROC analysis. We plan to investigate the effect of the matrix size of the coordinate system and other parameters on the performance of the neural network in order to optimize the selection of many of the parameters that are involved. When the performance of the neural network for correlating the locations in two views is optimized, we plan to apply the trained neural networks to indicate the range of potential locations of the lesion in a view by entering the location of a finding in another view. The range of potential locations of the lesion will be determined from the distribution of output values in units in the output layer. A polynomial surface-fitting technique will be used to obtain a smooth distribution from the distribution of output values.

In the detection of masses by comparing a current mammogram with a previous one, the previous mammogram will be selected from recent exams obtained within three years of the time the current mammogram is obtained. Prior to the subtraction of two mammograms, the previous mammogram will be preprocessed in order to gray-level match it to the current mammogram. This preprocessing step will involve gray level histogram modification and will only be performed if the difference in the

average gray level within the breast region is greater than a predetermined value such as 150 in terms of a 10-bit range. We will limit this preprocessing in this way since our preliminary data indicates that the temporal subtraction method is sufficiently robust to be independent of slight gray level variations.

Prior to subtracting the current and the previous mammograms, breast borders in the two mammograms need to be matched reasonably well. We plan to initially employ a technique for alignment of breast borders that we have successfully applied previously to match two mammograms of the right and left breasts (75), based on the rigid transformation of a coordinates system by shifting and rotation. In order to align the two breast images, landmarks are used to establish the correspondence between the two images. Breast borders (discussed earlier) and nipple positions are used as such landmarks. The nipple position in each breast image is identified using a technique we developed (75) on the basis of the thicker skin line and greater subcutaneous parenchymal opacity that is present around the nipple. Image registration is accomplished by use of a "constrained feature-matching" technique that involves two steps: (1) determining a constrained correspondence between points (x,y locations along the borders); and (2) matching the established corresponding points of one image with those of the other image by use of a least-squares method for translation and rotation. The least-squares method used to find the "best" match between the two sets (i.e., 2 matrices) of (x,y) locations corresponding to the border points employs the solution of the "Orthogonal Procrustes Problem" obtained by Schonemann et al. (90). In order to compensate for possible computer error in the identification of nipple positions, the identified nipple position of each breast image is varied within a certain range and the matching procedure (described earlier) is repeated for each variation. The best registration is determined by finding the smallest minimized sum of squares of the residual matrix from all possible nipple-position variations. Using the resulting optimal alignment parameters, the corrected previous breast image is translated and rotated relative to the current breast image. A common region, defined as the region enclosed by both the current border and aligned previous breast border, is then determined and used as a common border.

For detection of masses, smaller matrix images of approximately 512 x 512 for the current and previous mammograms will be first prepared by subsampling the digitized mammogram of approximately 2k x 2k matrix. This reduction of matrix size has been shown to be useful for detection of masses in our previous study, since mass sizes are generally larger than 5 mm and the pixel size of approximately 0.4 mm is adequate for detection of masses. We plan to investigate linear and nonlinear subtraction techniques for detection of masses by comparison of the current mammogram at each view with the corresponding previous mammogram. With linear subtraction, the two mammograms are subtracted and then gray-level thresholding is performed to segment the current image into potential locations of masses. With nonlinear subtraction (similar to that used in our bilateral subtractiontechnique for comparing left and right breast images), gray-level thresholding is performed prior to subtraction. This initial thresholding eliminates some normal anatomic background from further analysis. A selected number of images thresholded with different cut-off gray levels is obtained from the current mammogram. Subtraction of ten sets of corresponding current and previous mammograms, each thresholded at ten different levels determined from the histogram, is performed to generate ten temporal-subtraction images. A linking process then accumulates the information into an image, called a run-length image, where the value of each pixel in the image indicates how often the corresponding location in the set of 10 subtraction images have gray levels above a particular cut-off gray value. The run-length image is thresholded to yield the suspicious areas and submitted for feature extraction similar to that discussed in section (1)(c) except different features will be selected as well as the structure and weights of a neural networks in order to customize the method for elimination of false positives. Our preliminary studies, involving an evaluation on temporal cases from 76 patients, yielded a sensitivity of approximately 75% at 3 false-positive detections per image. Note that there was no preselection of cases except that these patients had prior films and that no optimization for elimination of false positives was performed (i.e., the neural network from the bilateral subtraction method was used). Thus, we believe that optimization of the subtraction method and the feature selection will increase sensitivity and specificity.

Results to date

We have evaluated the potential benefit of incorporating a temporal subtraction scheme with our bilateral subtraction technique for improving the sensitivity of mass detection (112). A database of 79 cases was used, each of which contained a lesion in at least the current exam. Two methods for image registration of the temporal images were investigated: one used translation and rotation based on computer-determined skin lines and the other used a warping technique based on the cross-correlation of regions of interest located throughout the parenchyma. The characteristics of the false-positive detection sresulting form the bilateral subtraction and teom temporal subtraction were analysed. The distribution of the true positives and false positives were simlar despite the fact that many of the false positives resulting from the two schemes were in different locations in the breast parenchyma. At a false-positive rate of four per image, the combined (Logical OR) scheme detected 85% of the masses, which was 8% greater than the bilateral subtraction technique alone. Although further work is needed to reduce the false-positive rate, the combined use of bilateral and temporal subtraction methods shows potential for an improvement in sensitivity in the detection of masses.

Incorporation of the computer-vision methods with an Mammo/Icon mammographic review system for enhanced diagnosis.

Experimental methods

Mammo/Icon has been designed to compliment a radiologists normal reading patterns. The system is used to report mammographic findings, using either speech recognition technology, or mouse interaction with a graphical user interface. As findings are recorded, Mammo/Icon functions in the background, collecting potentially helpful images from a knowledge base. In this proposal, we plan to modify the Mammo/Icon system so that the features of the lesion in question are automatically determined. The system would then automatically retrieve images of similar characteristics for display to the radiologist. Along with the images, the computer would also indicate an estimate of the likelihood of malignancy.

The Mammo/Icon mammography image review system contains a detailed lexicon of mammography image features and associated scalar values which allow stratification of images by pre-defined and user defined criteria. This scheme will be expanded to allow retrieval of images based on CAD derived features including the geometric-based, intensity-based, and gradient-based features, as well as ANN outputs. The Mammo/Icon database is separately being modified for compatibility with BIRADS nomenclature. Separate pre-defined search strategies are used to set up each axis of retrieved images. Additional heuristics determine the order that images are presented in each array. When the quantitative CAD derived parameters are added to the descriptors of each reference image in the Mammo/Icon database, additional heuristics will be added to the axis ordering strategies to arbitrate conflicts with the existing system.

Currently, Mammo/Icon contains nearly 400 images. All images will be submitted to the CAD system for computerized characterization. The values obtained for characterization of the lesions will be added to the image descriptors in the Mammo/Icon database.

We will implement CAD and Mammo/Icon systems on existing hardware at both Yale University and the University of Chicago. Initially we will need to develop an hardware interface between CAD and Mammo/Icon. When a new index case is processed by the CAD system, it will automatically pass derived data to the Mammo/Icon search engine and generate a search of the image database. We will develop a sequential database search. The initial search of the Mammo/Icon database will be generated solely by automatically extracted features by the CAD system. A second pass search will be generated by features described by the radiologist during case reporting. A third iteration will integrate steps one and two. To accommodate these complimentary methods, the Mammo/Icon user

interface will be modified to allow additional axis display (and display of characterization targets from CAD).

Results to date

Dr. Swetts at Yale is upgrading the program for the Mammo/Icon in order to incorporate more current mammographic reference images. Once his software component is complete it will be integrated into the University of Chicago classification scheme. However, due to the delay with the Yale software component, we have been preparing a system here at the UC that incorporates the computerized characterization of mass lesions and retrieves malignant and benign masses having similar features. The computer extracts the features described earlier for the computer classification. We are also accumulating a database of confirmed malignant and benign mass lesions (119 cancerous cases and 100 benign cases).

Evaluation of the CAD methods for mammography

Experimental methods

Performance studies will be done using a database of mammographic cases that have a distribution of subtle cases of normal, benign and malignant masses. "Truth" concerning the presence and malignancy of masses will be established with the aid of expert mammographers, follow-up reports and surgical biopsy reports. Normal cases will be selected from patients who have had normal follow-up exams. Prior films will be obtained for each patient.

Performance will be examined by calculating the fraction of lesions detected (true-positive rate) and the number of falsely-reported areas per case. The clinical database for the performance evaluations will ultimately contain 300 cases (100 normal, 50 with spiculated lesions, 50 with architectural distortions, 50 with small early cancers, and 50 with circumscribed lesions). The performance of the methods will be determined in terms of self-consistency results and round-robin (leave-one-out) or jackknife methods. We have used such methods many times in the past in our evaluation of other computerized schemes . ROC curves will be obtained by fitting continuous output data from the computerized scheme (such as number of pixels in the Hough spectrum domain above threshold for each ROI) using the LABROC4 program (87,97,98). The area under the ROC curve (A_z) will be used as an indicator of performance. Free-response ROC (FROC) analysis (99) and FROC-AFROC analysis (100) will be used in analyzing the data pertaining to localization of the abnormality. The ordinates of both FROC curves and AFROC curves are the fraction of lesions that are correctly localized by the observer. However, the abscissa of an FROC curve is the average number of false positives per image, whereas the abscissa of an AFROC curve is the probability of obtaining a false-positive image (i.e., an image containing one or more false-positive responses). We will fit an FROC curve to the performance data for each method being tested using Chakraborty's FROCFIT program (100). We have used this method in the past to fit mass detection data from our computerized scheme for the detection of masses using the bilateral subtraction approach and a single-image approach (68). The area A_1 under the AFROC curve, an alternative representation of the FROC curve, will be then used as the index with which to indicate performance.

Between the University of Chicago databases and Yale University Mammo/Icon database, we expect that 1000 radiographic lesions will be available for testing. Note that here the measure of performance will be the A_z value (from ROC analysis) obtained in the task of distinguishing between malignant and benign lesions.

Results to date

Databases are continuously being collected. For mass detection, we have approximately 119 clinical cases of malignant masses. New data for the classification database includes the 119 malignant cases as well as 100 benign cases. We also have additional 50 more malignant cases that have yet to be added to the database. The complete statistical evaluation will be performed at a later date when the databases are complete.

CONCLUSIONS

Computer-aided diagnosis has been successfully implemented in the clinical area for screening mammography using a prototype intelligent mammography workstation as a "second reader." The workstation provides for film digitization, image analysis, and display media for output of the computer aid. Introduction of the workstation did not result in an increase in callback rate. Results from the 3.5-year evaluation period are promising and plans are being made for a longer-term clinical study.

We are continuing to investigate methods for increasing the sensitivity and specificity of the computerized detection method. We have found that a combined bilateral-image and single-image method inproves the sensitivity for detection. We have also developed a new lesion segmentation method, which substantially improves lesion and feature extraction.

We have evaluated the potential benefit of incorporating a temporal subtraction scheme with our bilateral subtraction technique for improving the sensitivity of mass detection. Although further work is needed to reduce the false-positive rate, the combined use of bilateral and temporal subtraction methods shows potential for an improvement in sensitivity in the detection of masses.

We will also incorporate information from both the CC and MLO views into the detection task. We have already incorporated information from both the MLO and CC views in the computerized classification of masses. We have shown that the computer-extracted features and the computer decision making process yield a classification performance that is similar to that of an experienced mammographer. The computer classification uses a hybird system incorporating both a rule-based system and an artificial neural network. Our analysis of multiple classifiers showed that the hybrid system may be preferered when there exists a dominant lesion feature. We have also shown that our classification scheme is robust to film digitization using three different digitizers.

We are continuing to collect clinical databases for detection and classification. Upon completion of the database, statistical evaluation will be performed. In addition, these images are being incorporated into an automatic case retrival system for mass classification.

REFERENCES

1. Silverberg E, Boring CC, Squires TS: Cancer Statistics, 1990. CA 40: 9-27, 1990.
2. Tabar L, Dean PB: Basic principles of mammographic diagnosis. Diagn. Imag. Clin. Med. 54: 146-157, 1985.
3. American Cancer Society: CA Cancer J Clin 33: 255, 1983.
4. Baker L: CA Cancer J Clin 32: 194, 1982.
5. NCRP Report No. 85: Mammography (National Council on Radiation Protection: Washington, D.C., 1986).
6. Shapiro S, Venet WS, Strax PH, et al.: Ten to fourteen-year effect of screening on breast cancer mortality. JNCI 69: 349-355, 1982.
7. Verbeek ALM, Hendricks JH, Holland R, et al: Reduction of breast cancer mortality through mass screening with modern mammography. Lancet 1: 1222-1224, 1984.
8. Collette HJA, Day NE, et al.: Evaluation of screening for breast cancer in a non-randomized study (the DOM project) by means of a case-control study. Lancet 1: 1224, 1226, 1984.
9. Tabar L, Gad A, Holmberg LH, et al.: Reduction in mortality from breast cancer after mass screening with mammography. Randomized trial from the Breast Screening Working Groups of the Swedish National Board of Health and Welfare. Lancet 1: 829-832, 1985.
10. Andersson I, Aspegren L, Janzon L, et al.: Mammographic screening and nortality from breast cancer: The Malmo mammographic screening trial. Br. Med. J. 297:943,1988.
11. Feig SA: Decreased breast cancer mortality through mammographic screening: Results of clinical trials. Radiology 167: 659-665, 1988.
12. Black JW, Young B: A radiological and pathological study of the incidence of calcifications in diseases of the breast and neoplasms of other tissues. Br. J. Radiol. 58: 596-598, 1965.
13. Wolfe JN: Analysis of 462 breast carcinomas. AJR 121: 846-853, 1974.
14. Sickles EA: Mammographic detectability of breast microcalcifications. AJR 139: 913-918, 1982.
15. Fisher ER, Gregorio RM, Fisher B, et al.: The pathology of invasive breast cancer. Cancer 36: 1-84, 1975.
16. Millis RR, Davis R, Stacey AJ: The detection and significance of calcifications in the breast: A radiological and pathological study. Br. J. Radiol. 49: 12-26, 1976.
17. Murphy WA, DeSchryver-Kecskemeti K: Isolated clustered microcalcifications in the breast: Radiologic-pathologic correlation. Radiology 127: 335-341, 1978.
18. Muir BB, Lamb J, Anderson TJ: Microcalcification and its relationship to cancer of the breast: Experience in a screening clinic. Clin. Radiol. 149: 193-200, 1983.
19. Sickles EA: Mammographic features of 300 consecutive nonpalpable breast cancers. Am J Rad 146: 662-663, 1986.
20. Bassett LW, Gold RH: Breast cancer detection: Mammography and other methods in breast imaging, Grune and Stratton (New York), 1987.
21. Baines CJ, Miller AB, Wall C, McFarlane DV, et al.: Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: A preliminary report from five centers. Radiology 160: 295-298, 1986.
22. Pollei SR, Mettler FA, Bartow SA, Moradian G, Moskowitz M: Occult breast cancer: Prevalence and radiographic detectability. Radiology 163: 459-462, 1987.
23. Andersson I: What can we learn from interval carcinomas? Recent Results in Cancer Research 90: 161-163, 1984.
24. Martin JE, Moskowitz M, Milbrath JR: Breast cancers missed by mammography. AJR 132: 737, 1979.
25. Buchanann JR, Spratt JS, Heuser LS: Tumor growth, doubling times, and the inability of the radiologist to diagnose certain cancers. Radiologic Clinics of North America 21: 115, 1983.
26. Holland T, Mrvunac M, Hendriks JHCL, et al.: So-called interval cancers of the breast. Pathologic and radiographic analysis. Cancer 49:2527, 1982.
27. Murphy WA Jr, Destouet JM, Monsees BS: Professional quality assurance for mammography screening programs. Radiology 175: 319-320, 1990.

28. Bird RE: Professional quality assurance for mammography screening programs. Radiology 177: 587, 1990.
29. Brenner RJ: Medicolegal aspects of breast imaging: variable standards of care relating to different types of practice. AJR 156: 719-723, 1991.
30. Tabar L, Dean PB: Teaching Atlas of Mammography, George Thieme Verlag (Stuttgart, New York), 1983.
31. Moskowitz M: Screening for breast cancer: How effective are our tests? A critical review. Ca-A Cancer Journal for Clinicians 33: 26-39, 1983.
32. Giger ML: "Future of Breast Imaging. Computer-Aided Diagnosis". In:AAPM/RSNA Categorical Course on the Technical Aspects of Breast Imaging, 3rd edition, Eds. Haus A. and Yaffe M. 283-298, 1994.
33. Vyborny CJ, Giger ML: Computer vision and artificial intelligence in mammography. AJR 162: 699-708, 1994.
34. Winsberg F, Elkin M, Macy J, Bordaz V, Weymouth W: Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. Radiology 89: 211-215, 1967.
35. Spiesberger W: Mammogram inspection by computer. IEEE Transactions on Biomedical Engineering 26: 213-219, 1979.
36. Kimme C, O'Loughlin BJ, Sklansky J: Automatic detection of suspicious abnormalities in breast radiographs. In: Data Structures, Computer Graphics, and Pattern Recognition, edited by A. Klinger, K. S. Fu, T. L. Kunii (Academic Press, New York), 1975, pp. 427-447.
37. Hand W, Semmlow JL, Ackerman LV, Alcorn FS: Computer screening of xeromammograms: A technique for defining suspicious areas of the breast. Computers and Biomedical Research 12: 445-460, 1979.
38. Semmlow JL, Shadagoppan A, Ackerman LV, Hand W, Alcorn FS: A fully automated system for screening xeromammograms. Computers and Biomedical Research 13: 350-362, 1980.
39. Ackerman LV, Gose EE: Breast lesion classification by computer and xeroradiography. Cancer 30: 1025-1035, 1972.
40. Ackerman LV, Mucciardi AN, Gose EE, Alcorn FS: Classification of benign and malignant breast tumors on the basis of 36 radiographic properties. Cancer 31: 342-352, 1973.
41. Wee WG, Moskowitz M, Chang N-C, Ting Y-C, Pemmeraju S: Evaluation of mammographic calcifications using a computer program. Radiology 116:717-720, 1975.
42. Fox SH, Pujare UM, Wee WG, Moskowitz M, Hutter RVP: A computer analysis of mammographic microcalcifications: Global approach. Proc. IEEE 5th International Conf. on Pattern Recognition: 624-631, 1980.
43. Magnin IE, Cluzeau F, Odet CL, Bremond A: Mammographic texture analysis: an evaluation of risk for developing breast cancer. Optical Engineering 25: 780-784, 1986.
44. Caldwell CB, Stapleton SJ, Holdsworth DW, Jong RA, Weiser WJ, Cooke G, Yaffe MJ: Characterization of mammographic parenchymal pattern by fractal dimensions. Phys. Med. Biol. 35: 235-247, 1990.
45. Fam BW, Olson SL, Winter PF, Scholz FJ: Algorithm for the detection of fine clustered calcifications on film mammograms. Radiology 169: 333-337, 1988.
46. Olson SL, Fam BW, Winter PF, Scholz FJ, Lee AK, Gordon SE: Breast calcifications: Analysis of imaging properties. Radiology 169: 329-331, 1988.
47. Davies DH, Dance DR: Automatic computer detection of clustered calcifications in digital mammograms. Phys. Med. Biol. 35: 111-118, 1990.
48. Astley S, Hutt I, Adamson S, et al. Automation in mammography: computer vision and human perception. Proc SPIE 1993;1905:716-730
49. Grimaud M, Muller S, Meyer F: Automated detection of microcalcifications in mammograms. Radiology 177(P):288, 1990.
50. Mascio LN, Hernandez JM, Logan CM. Automated analysis for microcalcifications in high resolution digital mammograms. Proc SPIE 1898: 1993:(in press).

51. Karssemeijer N: A stochastic method for automated detection of microcalcifications in digital mammograms. Information Processing in Medical Imaging, Springer-Verlag (New York), pp.227-238, 1991.
52. Lai SM, Li X, Bischof WF: On techniques for detecting circumscribed masses in mammograms. IEEE Transactions on Medical Imaging 8: 377-386, 1989.
53. Brzakovic D, Luo XM, Brzakovic P: An approach to automated detection of tumors in mammograms. IEEE Transactions on Medical Imaging 9: 233-241, 1990.
54. Kegelmeyer WP: Computer detection of stellate lesions in mammograms. Proc. SPIE 1660: 446-454, 1992.
55. Gale AG, Roebuck EJ, Riley P, Worthington BS, et al.: Computer aids to mammographic diagnosis. British Journal of Radiology 60: 887-891, 1987.
56. Getty DJ, Pickett RM, D'Orsi CJ, Swets JA: Enhanced interpretation of diagnostic images. Invest. Radiol. 23: 240-252, 1988.
57. Swett HA, Miller PA: ICON: A computer-based approach to differential diagnosis in radiology. Radiology 163: 555-558, 1987.
58. Swett HA, Fisher PR, Cohn AI, Miller PI, Mutalik PG: Expert system controlled image display. Radiology 172: 487-493, 1989.
59. Chan HP, Doi K, Galhotra S, Vyborny CJ, MacMahon H, Jokich PM: Image feature analysis and computer-aided diagnosis in digital radiography. 1. Automated detection of microcalcifications in mammography. Med Phys 14: 538-548, 1987.
60. Chan HP, Doi K, Vyborny CJ, Lam KL, Schmidt RA: Computer-aided detection of microcalcifications in mammograms: Methodology and preliminary clinical study. Invest Radiol 23: 664-671, 1988.
61. Chan HP, Doi K, Vyborny CJ, Schmidt RA, Metz CE, Lam KL, Ogura T, Wu Y, MacMahon H: Improvement in radiologists' detection of clustered microcalcifications on mammograms: The Potential of computer-aided diagnosis. Invest Radiol 25: 1102-1110, 1990.
62. Nishikawa RM, Doi K, Giger ML, Yoshimura H, Wu Y, Vyborny CJ, Schmidt RA, Chan HP: Use of morphological filters in the computerized detection of microcalcifications in digital mammograms. Medical Physics 17: 524, 1990.
63. Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computer-aided detection of clustered microcalcifications: An improved method for grouping detected signals. Med Phys 20: 1661-1666, 1993.
64. Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE, Wu Y, et al.: Computer-aided detection and diagnosis of masses and clustered microcalcifications from digital mammograms. Proc. SPIE 1905: 422-432, 1993.
65. Wu Y, Doi K, Giger ML, Nishikawa RM: Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks. Radiology 19: 555-560, 1992.
66. Giger ML, Yin F-F, Doi K, Metz CE, Schmidt RA, Vyborny CJ: Investigation of methods for the computerized detection and analysis of mammographic masses. Proc. SPIE 1233: 183-184, 1990.
67. Yin F-F, Giger ML, Doi K, Metz CE, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Analysis of bilateral-subtraction images. Medical Physics 18: 955-963, 1991.
68. Yin FF, Giger ML, Vyborny CJ, Doi K, Schmidt RA: Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. Invest Radiol 28: 473-481, 1993.
69. Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE: Application of neural networks in mammography: Applications in decision making in the diagnosis of breast cancer. Radiology 187: 81-87, 1993.
70. Giger ML, Nishikawa RM, Doi K, Yin FF, Vyborny CJ, Schmidt RA, Metz CE, Wu Y, MacMahon H, Yoshimura H: Development of a "smart" workstation for use in mammography. Proc. SPIE 1445: 101-103, 1991.

71. Doi K, Giger ML, MacMahon H, Hoffmann KR, et al.: Computer-aided diagnosis: development of automated schemes for quantitative analysis of radiographic images. Seminars in Ultrasound, CT and MR 13: 140-152, 1992.
72. Giger ML, Doi K, MacMahon H, Nishikawa RM, Hofmann KR, et al.: An "intelligent" workstation for computer-aided diagnosis". RadioGraphics 13: 647-656, 1993.
73. Nishikawa RM, Giger ML, Doi K, Vyborny CJ: Effect of case selection on the performance of computer-aided detection schemes. Med Phys 21:265-269, 1994..
74. Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Investigation of feature-analysis techniques. Journal of Digital Imaging 7:18-26, 1994.
75. Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. Med Phys 21:445-452, 1994.
76. Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computer-aided detection of clustered microcalcifications on digital mammograms. Medical and Biological Engineering and Computing 33:174-178, 1995.
77. Giger ML, Vyborny CJ, Schmidt RA: Computerized characterization of mammographic masses: Analysis of spiculation. Cancer Letters 77: 201-211, 1994.
78. Bick U, Giger ML, Huo Z, Schmidt RA, Doi K, Nishikawa RM, Vyborny CJ: Automated detection of skin thickening in mammograms. Proc. CAR '93 pgs. 461-465, 1993.
79. Bick U, Giger ML, Schmidt RA, Nishikawa RM, Wolverton DE, Lu P, Vyborny CJ, Doi K: Automated segmentation of digitized mammograms. Academic Radiology 2: 1-9, 1995.
80. Jiang Y, Nishikawa RM, Giger ML, Doi K, Schmidt RA, Bick U, et al.: Analysis of image features for automated classification of malignant and benign clustered microcalcifications. Radiology 189(P): 317, 1993.
81. Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA: Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. Med Phys 21:517-524, 1994.
82. Huo Z, Giger ML, Vyborny CJ, Bick U, Lu P, Wolverton DE, Schmidt RA: Analysis of spiculation in the computerized classification of mammographic masses" Medical Physics 22:1569-1579, 1995.
83. Bick U, Giger ML, et al: A new single-image method for computer-aided detection of small mammographic masses. Proc. CAR '95, LemkeHU, Inamura K, Jaffe CC, Vannier MW, eds. pgs. 357-363, 1995.
84. Zhang M, Giger ML: Automated detection of spiculated lesions and architectural distortions in digitized mammograms. Proc SPIE 2434: 846-854, 1995.
85. Nishikawa RM, Haldemann RC, Papaioannou J, Giger ML, Lu P, Schmidt RA, Wolverton DE, Bick U, Doi K: Initial experience with a prototype clinical "intelligent" mammography workstation for computer-aided diagnosis. Proc SPIE 2434 (in press), 1995.
86. Kupinski M, Giger ML, Lu P, Huo Z: Computerized detection of mammographic lesions: Performance of artificial neural network with enhanced feature extraction. Proc SPIE 2434 (in press), 1995.
87. Metz CE: ROC methodology in radiologic imaging. Invest Radiol 21: 720-733, 1986.
88. Matsumoto T, Yoshimura H, Doi K, Giger ML, Kano A, MacMahon H, Abe K, Montner SM: Image feature analysis of false-positive diagnoses produced by automated detection of lung nodules. Invest Radiol 27: 587-597, 1992.
89. Giger ML, Doi K: Investigation of basic imaging properties in digital radiography. 3. Effect of Pixel Size on SNR and Threshold Contrast. Med. Phys. 12: 201-208, 1985.
90. Giger ML, Doi K: Effect of pixel size on detectability of low-contrast signals in digital radiography. J of the Optical Society of America A 4: 966-975, 1987.
91. Ohara K, Doi K, Metz CE, Giger ML: Investigation of basic imaging properties in digital radiography. 13. Effect of structured noise on the detectability of simulated stenotic lesions. Med Phys 16:14-21, 1989.

92. Ohara K, Chan HP, Doi K, Giger ML, Fujita H: Investigation of basic imaging properties in digital radiography. 8. Detection of simulated low-contrast objects in DSA images. Med Phys 13: 304-311, 1986.
93. MacMahon H, Metz CE, Doi K, Kim T, Giger ML, Chan H-P: The effect of display format on diagnostic accuracy in digital chest radiography: A comparison of hardcopy, video, and reversed grey scale. Radiology 168: 669-673, 1988.
94. MacMahon H, Doi K, Sanada S, Montner SM, Giger ML, et al.: Effect of data compression on diagnostic accuracy in digital chest radiography. An ROC study. Radiology 178: 175-179, 1991.
95. MacMahon H, Sanada S, Doi K, Giger ML, Xu X-W, Yin F-F, Montner SM, Carlin M: Direct comparison of conventional and computed radiography with a dual image recording technique. RadioGraphics 11: 259-268, 1991.
96. Giger ML, Doi K, MacMahon H, Metz CE, Yin FF: Computer-aided detection of pulmonary nodules in digital chest images. RadioGraphics 10: 41-52, 1990.
97. Dorfman DD, Alf E: Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals-rating method data. J Math Psych 6: 487-498, 1969.
98. Metz CE: Some practical issues of experimental design and data analysis in radiological ROC studies. Invest. Radiol. 24: 234-245, 1989.
99. Bunch PC, Hamilton JF, et al.: A free response approach to the measurement and characterization of radiographic observer performance. Proc. SPIE 127: 124-135, 1977.
100. Chakraborty D: Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data. Med. Phys. 16: 561-568, 1989.
101. Bamber D. The area above the ordinal dominance graph and the area below the receiver operating graph. J. Math Psych 12: 387-415, 1975.
102. Kupinski M. Computer-aided diagnosis of mammographic lesions: Genetic algorithms in the optimizationof artificial neural network inputs.Thesis, Trinity University/The University of Chicago.
103. Schonemann. A generalized solution of the orthogonal procrustes problem. Psychometrika 31:1-10, 1966.
104. Bick U, Giger ML, Schmidt RA, Nishikawa RM, Wolverton DE, Doi K: Computer-aided breast cancer detection in screening mammography. Digital Mammography '96. Proceedings of the 3rd International Workshop of Digital Mammography, Elsevier, New York, pp. 97-104, 1996.
105. Zhang M, Giger ML, Vyborny CJ, Doi K: Mammographic texture analysis for the detectio of spiculated lesions. Digital Mammography '96. Proceedings of the 3rd International Workshop of Digital Mammography, Elsevier, New York, pp. 347-350, 1996.
106. Nishikawa RA, Giger ML, Wolverton DE, et al: Prospective testing of a clinical mammography workstation for CAD: Analysis of the first 10,000 cases. Digital Mammography '98. Proceedings of the 4th International Workshop of Digital Mammography, The Netherlands, (in press), 1998.
107. Kupinski MA, Giger ML: Automated seeded lesion segmentation on digital mammograms. IEEE Trans. on Medical Imaging (in press), 1998.
108. Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Schmidt RA, Doi K: Automated computerized classification of malignant and benign mass lesions on digitized mammograms. Academic Radiology 5: 155-168, 1998
109. Huo Z, Giger ML, Metz CE: Computerized classification of mammographic lesions: Effect of dominant features on neural network performance. Medical Physics, (in review), 1998.
110. Jiang Y, Metz CE and Nishikawa RM. A receiver operating characteristics partial area index for highly sensitive diagnostic tests. Radiology 1996; 201:745-750.
111. Metz CE, Herman BA and Shen J. Maximum-likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Statistics in Medicine (in press).
112. Zouras WK, Giger ML, Lu P, Wolverton DE, Vyborny CJ, Doi K: Investigation of a temporal subtraction scheme for computerized detection of breast masses in mammograms. Digital Mammography '96. Proceedings of the 3rd International Workshop of Digital Mammography, Elsevier, New York, pp. 411-416, 1996.